

Several multiplexes in the same city: The role of socioeconomic differences in urban mobility

Laura Lotero, Alessio Cadillo, Rafael Hurtado and Jesús Gómez-Gardeñes

Abstract In this work we analyze the architecture of real urban mobility networks from the multiplex perspective. In particular, based on empirical data about the mobility patterns in the cities of Bogotá and Medellín, each city is represented by six multiplex networks, each one representing the origin-destination trips performed by a subset of the population corresponding to a particular socioeconomic status. The nodes of each multiplex are the different urban locations whereas links represent the existence of a trip from one node (origin) to another (destination). On the other hand, the different layers of each multiplex correspond to the different existing transportation modes. By exploiting the characterization of multiplex transportation networks combining different transportation modes, we aim at characterizing the mobility patterns of each subset of the population. Our results show that the socioeconomic characteristics of the population have an extraordinary impact in the layer organization of these multiplex systems.

Laura Lotero

Departamento de Ciencias de la Computación y de la Decisión, Universidad Nacional de Colombia, Medellín, Colombia, e-mail: llotero0@unal.edu.co

Alessio Cardillo

Laboratoire de Biophysique Statistique, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland. Instituto de Biocomputación y Física de Sistemas Complejos, Universidad de Zaragoza, E-50018 Zaragoza, Spain, e-mail: alessio.cardillo@epfl.ch

Rafael Hurtado

Departamento de Física, Universidad Nacional de Colombia, Bogotá, Colombia, e-mail: rghurtadoh@unal.edu.co

Jesús Gómez-Gardeñes

Departamento de Física de la Materia Condensada, Universidad de Zaragoza, E-50009 Zaragoza, Spain. Instituto de Biocomputación y Física de Sistemas Complejos, Universidad de Zaragoza, E-50018 Zaragoza, Spain, e-mail: gardenes@gmail.com

1 Introduction

Understanding human mobility patterns have attracted, for decades, the attention of researchers from many different scientific realms. The first models based on empirical observations date back to the 40's and were elaborated by the sociologist Samuel A. Stauffer [1] and the philologist George K. Zipf [2]. It is at the end of the last century when, with the consolidation of mathematical frameworks such as the well-known *Gravity model* [3], when transportation science, as a realm of operations research, became a discipline on its own.

The advent of the big data era, have spurred the activity on transportation science and provided detailed datasets of real transportation systems. This characterization spans across many scales, from the short-range mobility patterns in urban areas [4, 5, 6] to world wide trips [7]. Remarkably, different degrees of resolution and types of information are nowadays available from the combined use of techniques for data gathering [8]. From the traditional datasets based on direct surveys [9], allowing to know the purpose of the trip (work/school, leisure, etc), to those large-scale ones gathered by tracking mobile communication systems [10, 11] or transport electronic cards [12]. This burst of activity have attracted many scientist from theoretical disciplines to contribute to the subject through the formulation of mobility models and mathematical tools aimed at reproducing and characterizing the observed patterns of movement [13, 14, 15, 16].

The rapid change in the patterns of human mobility in the last decades, specially in what concerns the decrease in their duration together with the increase of their length, makes its characterization of utmost importance for many disciplines beyond the traditional scope of transportation science. The most paradigmatic example is the relevance of human mobility in the spread of diseases. The inclusion of the mobility ingredient into epidemic models has allowed to design sophisticated theoretical frameworks aiming at forecasting the onset and duration of pandemics with high time and spatial resolution [17, 18, 19, 20, 21, 22].

In the last fifteen years, networks science [23, 24, 25] have appeared as the best suited mathematical frameworks to accommodate and characterize the interaction backbone of the very many complex systems captured by big data techniques. In fact, complex networks had been proposed as the natural framework to study spatially embedded systems [26] and, in particular, mobility networks. In these networks the different origins and destinations are represented as nodes of a graph, whereas the movements between locations are encoded as links connecting them [27]. Recently, thanks to the availability of more detailed information, it has been possible to represent many different types of transportation modes used for the movements within the same area under multilayer networks [28, 29, 30], in which network layer represents a single transportation mode. In this way, each node still represents a particular origin/destination location and it is present in each of the network layers. However, links are represented in a different layer of interaction depending on the kind of transportation mode used for connecting two locations. This particular multilayer network is usually termed as *multiplex*.

In the recent years, different human mobility systems have been addressed under the paradigm of multiplex networks, ranging from urban movements [31] to medium [32] and large scale trips [33, 34]. Following this approach, here we address the multiplex structure of urban mobility in two different cities: Bogotá and Medellín. The novelty of the results presented rely on an additional ingredient of the mobility patterns that, up to our knowledge, has been ignored up to date. This new ingredient is the socioeconomic status (*SES*) of the individuals, mainly related to their wealth. Being the composition of many cities in the world highly hierarchical and inhomogeneous in terms of the capital distribution, it is thus relevant to unveil the influence that the different SES have on the mobility patterns.

To this aim, and considering that another relevant ingredient included in the available data sets is the transportation modes used by the individuals, we analyze the mobility patterns in terms of a multiplex network. In particular, we will analyze six different multiplex networks, each one corresponding to a different SES. Our approach relies on the *adiabatic projection* technique, introduced in [34], that consist in monitoring how the structural properties of the aggregate network show up as a result of the merging of the layers composing the multiplex. Thanks to this approach, it has been possible to spotlight how *segregation* and *multimodality* are characteristic of some particular social classes, and to unveil the dominant role played by the middle-class in the utilization of the transportation system as a whole.

The structure of the paper is the following. We will first introduce the datasets used in section 2 and the adiabatic projection technique together with the topological estimators in which it is used in section 3. Section 4 is devoted to present the results of applying the former technique to the datasets of the cities of Bogotá and Medellín. Finally, in section 5 we draw some conclusions and future work perspectives.

2 Urban Mobility and Socioeconomic Status

The mobility data presented and analyzed here are taken from surveys carried out in two major cities of Colombia: Bogotá and Medellín. These surveys were originally designed to collect information about travelers and their trips, so to identify traffic patterns and apply the results to urban and transportation planning. In these surveys, each householder is asked about the trips performed the day before the interview, providing with the origin and destination zones, the departure and arrival times, the transportation mode used and the purpose of each trip. In addition, householders are characterized by their socioeconomic characteristics, such as the age, gender, occupation, and the socioeconomic characteristic of their housing, which it is defined as its SES. The survey for the city of Bogotá, having a population of about 7 million of inhabitants, has a sample size of 45446 people interviewed, reporting 100846 trips [40]. On the other hand, the survey for the metropolitan area of Medellín, with a population of about 3.5 million people, reports 127849 trips from 56513 personal interviews [41]. However, not all the people interviewed made a trip and thus the number of travelers in both cities is smaller (see Table 1).

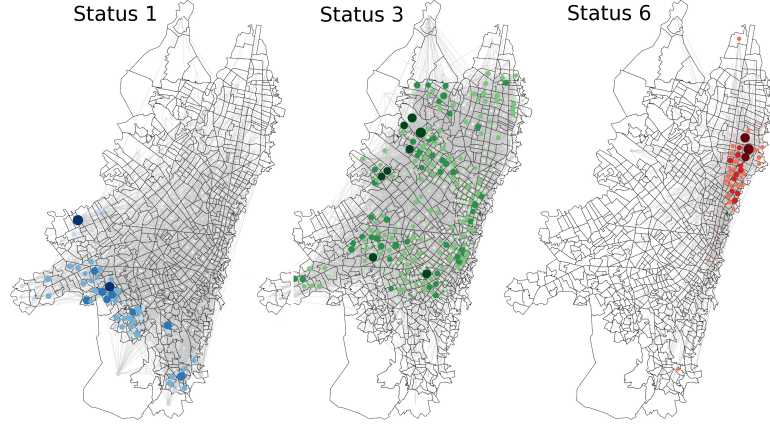


Fig. 1 Maps of the trips made in the city of Bogotá. Each map refers to a particular socioeconomic status, namely (from left to right) 1, 3 and 6. Each node corresponds to a different urban mobility zone, while a link (light gray) between two nodes indicates that a displacement has occurred among them. The size of the nodes indicates the amount of displacements (both from and to) occurring in that zone.

In Table 1 we briefly show the most relevant information about the population interviewed in both cities. From the network perspective, the mobility graphs derived from these surveys contain $N = 912$ nodes (being both origins and destinations) for the city of Bogotá, and $N = 413$ for the metropolitan area of Medellín. In this way, two nodes are linked whenever the survey reports the existence of at least one trip between two zones. In addition, we take advantage of the socioeconomic information provided by the surveys, in particular the information about the SES of each individual, being this a good proxy of the population wealth. This categorization of the population into strata is specific of Colombia, and ranges from status 1 for the lowest-income householders up to 6 for the highest-income individuals. Examples of mobility graphs of three of these socioeconomic groups are displayed in Figure 1.

As introduced above, the aim of our work is to study the different means of transportations coexisting in the urban mobility as a multiplex mobility network. Since the surveys contained a number of different transportation means (25 in Bogotá and 17 in Medellín) we grouped these transportation modes into 6 different categories. In particular: (i) pedestrian (walking), (ii) public transport, (iii) private transport

	P	$\langle A \rangle$	$Fem(\%)$	T_{TOT}	$\langle T \rangle$	$\langle n_s \rangle$	N	E
Bogotá	37483	33.23	54.4	100846	2.69	—	912	24588
Medellín	45496	33.34	53.4	127849	1.58	1.105273	413	18442

Table 1 Supplementary information about the mobility interviews. From left to right: number of travellers, P , their average age, $\langle A \rangle$, percentage of female subjects, Fem , total number of trips recorded, T_{TOT} , average number of trips per person, $\langle T \rangle$, average number of steps per trip, $\langle n_s \rangle$, number of urban areas (nodes), N , number of connections between areas (links), E .

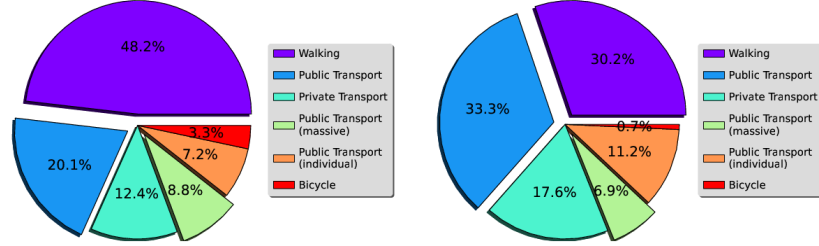


Fig. 2 Usage of the six transportation modes for the case of Bogotá (left) and Medellín (right).

(*e.g.* car or motorbike), (*iv*) public massive transport (*e.g.* metro), (*v*) public individual transport (*e.g.* taxi), and (*vi*) bicycle. The usage of each transportation group is displayed in Figure 2. Surprisingly, being the same classification for both datasets, we notice that the usage of each group is not the same in both cities. This is due to several factors, such as the different morphology of the cities and the differences in their urban development and planning. Regarding the socioeconomic composition of the population we report in Figure 3 the partition of the samples used in the surveys of Bogotá and Medellín in agreement with the real socioeconomic distribution of both cities, being the majority of the population in SES 2 and 3.

Our goal in the next sections is to use the mobility and socioeconomic data provided by these surveys, to explain how the SES of individuals affect the mobility patterns. To illustrate how the different SES make use of the available transportation modes we show in Figure 4 two different mobility matrices, \mathbf{M} . The first type of matrix (top) shows how the usage of a transportation mean is partitioned into the different strata, whereas the second one shows how individuals of a particular SES use the different available modes. In both cases, the latter information is casted in 6×6 matrices which entries $M_{t,s}$ corresponds to, in the first case, the fraction of trips that individuals belonging to SES s perform using transportation mode t . In its turn, in the second case, the entry $M_{t,s}$ accounts for the probability that a trip of an individual of SES s is performed using mean t . This information is shown for both cities, Bogotá (left) and Medellín (right).

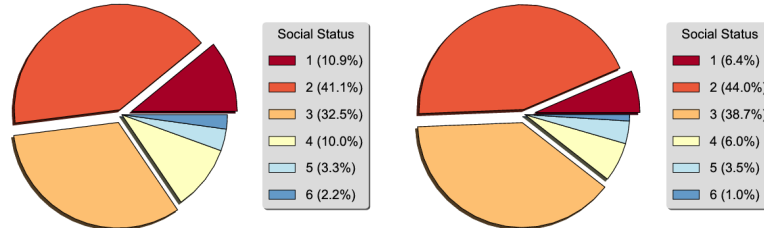


Fig. 3 Socioeconomic distribution of the population for the cities of Bogotá (left) and Medellín (right). SES ranging from 1 (poorest) to 6 (richest) is assigned to every traveler according to his/her economic wealth.

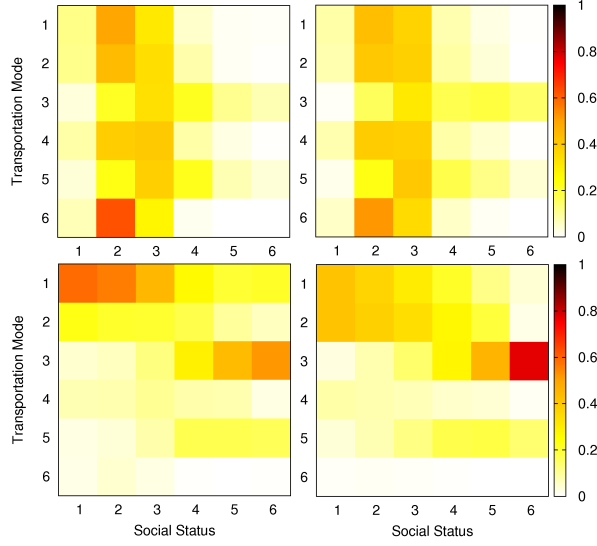


Fig. 4 Distribution of usage of transportation modes by the 6 SES. The matrices in the left (right) correspond to the city of Bogotá (Medellín). The matrices in the top (bottom) are normalized as $\sum_s M_{ts} = 1$ ($\sum_t M_{ts} = 1$). Transportation modes, from 1 to 6, are ordered according to Figure 2.

At first sight, comparing those matrices in the left (Bogotá) and those in the right (Medellín), both cities display roughly the same usage patterns. In particular, concerning the matrices in the top we can observe that the usage of modes follows two different patterns depending on the precise transportation mean. For instance, modes 3 (private transport) and 5 (public individual transport) accumulate users from SES 2 to 6, whereas modes 4 (public massive transport) and 6 (bicycle) are mainly used by individuals with SES 2 and 3.

The second group of matrices (bottom) show, both for Bogotá and Medellín, that although multimodality is somehow present for individuals from SES 3 and 4, there is a high tendency to concentrate the trips around few transportation means. In fact, this concentration makes clear the socioeconomic differences according to the means selected: poor individuals from SES 1 and 2 concentrate their trips using modes 1 (pedestrian) and 2 (public transport), which are the cheapest ones, while those belonging to SES 5 and 6, mostly use means 3 (private transport) and 5 (public individual transport), that represent the most expensive ones. Thus, in the first case the concentration of trips around means 1 and 2 is due to the segregation of SES 1 and 2 towards cheap means whereas individuals from SES 5 and 6 can select their means according to their commodity. Thus, the mobility patterns in both cities show a clear transition segregation-multimodality-selection when going from the poorest to the richest.

3 The Adiabatic Projection of a Multiplex

From the analysis of the mobility matrices in Figure 4, it becomes clear that the usage of the different transportation modes depends strongly of the status of the individuals. These results demand the analysis of how the different transportation modes are associated forming a *mobility multiplex network* (MMN) for each social status. In this section we present the *Adiabatic Projection* (AP) technique used to characterize MMN and the structural quantities under study.

Following the formalism introduced by Battiston *et al.* [39], we consider the MMN of a social status s as a system composed of N nodes and $M = 6$ layers. As explained before, nodes correspond to the different urban areas in a city. Layers, instead, represent different transportation modes. Keeping in mind such setup, and particularizing in the mobility multiplex of a given social status s , it is possible to associate to each layer α ($\alpha = 1, \dots, M$) a graph $\mathcal{G}^{[s,\alpha]}(N, \mathcal{E}^{[s,\alpha]})$ described by an adjacency matrix $\mathbf{A}^{[s,\alpha]}$ whose entries are defined as $a_{ij}^{[s,\alpha]} = 1$ if zones i and j are connected by (at least) a trip of an individual from status s using transportation mode α . Under this formalism, the MMN of a social status s is fully described by the so-called *vector of adjacency matrices* \mathbf{A}^s given by:

$$\mathbf{A}^s = \left\{ \mathbf{A}^{[s,1]}, \dots, \mathbf{A}^{[s,M]} \right\}. \quad (1)$$

Once having introduced the basic notation characterizing each of the MMN, we describe the AP procedure used to study the coexistence of several interaction (here transportation) modes in a multiplex network. The technique relies in merging together a subset $V(m)$ containing $m \leq M$ layers into a single (monolayer) graph $\mathcal{G}^{[s,V(m)]}(N, \mathcal{E}^{[s,V(m)]})$ where:

$$\mathcal{E}^{[s,V(m)]} = \bigcup_{\alpha \in V(m)} \mathcal{E}^{[s,\alpha]}. \quad (2)$$

Therefore, the network $\mathcal{G}^{[s,V(m)]}$ is obtained by projecting all the layers contained in $V(m)$ onto a single one and by converting the multiple links (those existing in several layers in $V(m)$) into single ones. In this way, the topology of the resulting projected network is described by the *projected adjacency matrix* $\mathbf{A}^{[s,V(m)]}$ defined as:

$$a_{ij}^{[s,V(m)]} = \begin{cases} 1 & \text{if } \sum_{\alpha \in V(m)} a_{ij}^{[s,\alpha]} > 0 \\ 0 & \text{if } \sum_{\alpha \in V(m)} a_{ij}^{[s,\alpha]} = 0. \end{cases} \quad (3)$$

The purpose of the AP of the layers of a multiplex is to analyze the evolution of some topological quantities when passing from single layers to the projected network resulting from merging all the M layers of the multiplex. Thus, the approach, introduced in [34] to study the *European Air Transportation Multiplex*, consists in varying the number of layers contained in the subset $V(m)$ from $m = 1$ to $m = M$. It is important to notice that the AP method (as introduced in [34]) considers, for each value m , the set $\mathcal{V}(m)$ containing all the possible subsets $V(m)$ comprising m

layers. In this way, given a topological quantity x , one evaluates x in each projected graph $\mathcal{G}^{[s,V(m)]}$ derived from each subset $V(m)$ contained in $\mathcal{V}(m)$ and average the values obtained over all the resulting graph. Thus, given m , the average value of x in $\mathcal{V}(m)$ reads:

$$\langle x \rangle(m) = \frac{m!(M-m)!}{M!} \sum_{V(m) \in \mathcal{V}(m)} x(\mathcal{G}^{[s,V(m)]}) . \quad (4)$$

Note that, although for $m = 1$ there are M possible subsets in $\mathcal{V}(1)$ whereas for $m = M$ there is only 1 subset $V(M)$, for a general value m the cardinal of $\mathcal{V}(M)$ can be extremely large. Thus, the AP technique demands a computationally expensive statistical treatment to cover all the possible layer combinations included in the sum of equation 4 when the number of layers M is large enough.

Here, instead, we get rid off the statistics over the sets $\mathcal{V}(m)$. In particular, based on the details contained in the dataset, we make use of the information about the usage of each transportation mode α by each SES s so that, for a certain value of m , we consider the projected graph $\mathcal{G}^{[s,V(m)]}$ constructed by merging the m most used transportation modes (layers) by SES s . In this way, for each value of m , $\mathcal{V}(m)$ contains one single subset $V(m)$ and thus we will denote each projected graphs as $\mathcal{G}^{[s,m]}$ and its associated adjacency matrix as $\mathbf{A}^{[s,m]}$. Apart from the computational simplification of this variant of the AP technique, the new path from $m = 1$ to $m = M$ informs about how the individuals of a particular SES are benefited by adding transportation modes to their trips allowing to distinguish between strata displaying either segregation or selection of modes and those socioeconomic compartments showing multimodality.

The topological quantities studied under the AP technique cover traditional structural measures, used in simple networks, and others that take into account the layer structure of a multiplex. In particular, for each graph $\mathcal{G}^{[s,m]}$ we will study the following usual properties:

- *The size, S , of the giant component and the number of components, n_c .* It is important to note that S is normalized to be $0 \leq S \leq 1$, so that $S = 1$ when the N nodes in the network take part of a unique component. In addition, to compute n_c we have considered that isolated nodes do not constitute a component so that components contributing to n_c are those of size equal or larger than 2.
- *The average path length, L .* As usual, L is the average length of the shortest paths among all the couples of nodes in the network. Since the networks under study are highly disconnected, especially for small values of m , we have adopted the typical way out to avoid divergences in L , *i.e.*, to consider only the nodes in the giant component.
- *The average degree, $\langle k \rangle$.* Again, in order to compute the average number of connections of the nodes we have excluded isolated nodes.
- *The clustering coefficient, C .* As usual, the clustering coefficient shows the probability that two nodes i and j having a common neighbor l are also connected. In this case also, isolated nodes do not contribute to clustering.

The above measures are those traditionally used for characterizing simple (single-layer) networks. However, there also exist measures that are specifically designed for multiplex networks (see the recent reviews [29, 30]). This is the case of the *Overlap*, O . The overlap quantifies the redundancy of links between layers, *i.e.*, the fact that a link between two given nodes i and j is present in several layers. In our multiplex networks the existence of a large overlap would imply a large tendency of the individuals (belonging to the same SES) of using different means of transportation for connecting the same urban areas i and j . In the recent years, several overlap measures have been proposed [35, 36, 37, 38, 39]. Here, for a given value of the number of projected layers m , the overlap of the resulting graph $\mathcal{G}^{[s,m]}$ is measured via two different quantities, namely:

$$O_1 = \frac{W - K}{K}, \quad (5)$$

$$O_2 = \frac{D}{K}. \quad (6)$$

Where K is the number of links in the aggregate graph $\mathcal{G}^{[s,m]}$, W is the total sum of the links in each of the m layers merged in $\mathcal{G}^{[s,m]}$, and D is the number of redundant links in the set of m layers. We can express these quantities, W , K and D , making use of the adjacency matrices associated to each layer, $\mathcal{G}^{[s,\alpha]}$, and that of the projection of the m most used layers, $\mathcal{G}^{[s,m]}$, as:

$$W = \sum_{\alpha=1}^m \sum_{i,j=1}^N a_{ij}^{s,\alpha}, \quad (7)$$

$$K = \sum_{i,j=1}^N a_{ij}^{s,m}, \quad (8)$$

$$D = \Theta \left(\sum_{\alpha=1}^m a_{i,j}^{s,\alpha} - 2 \right), \quad (9)$$

where, in the last equation, $\Theta(z)$ is the step function defined as $\Theta(z) = 0$ for $z < 0$ and $\Theta(z) = 1$ otherwise.

4 Results

In this section, and relying on the the AP technique of the MMN, our aim is to unveil the mobility patterns associated to the use of the transportation modes of each SES and, moreover, to monitor how the different patterns present in the transportation layers are combined into their corresponding mobility networks.

We start by studying how the combination of different transportation modes cover the different urban areas. This view can be explained as a percolation process driven by the addition of network layers (instead of nodes or links as in the

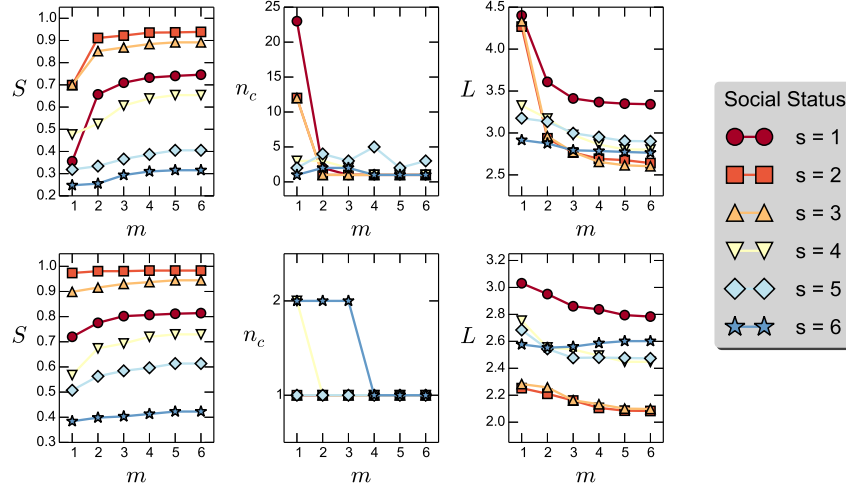


Fig. 5 Adiabatic evolution of the size of the giant component $S(m)$ (left panels), number of components $n_c(m)$ (central panels), and average path length $L(m)$ (right panels). The top panels are for the city of Bogotá, while those in the bottom refer to Medellín. Each curve corresponds to a different SES. Namely: 1 (●), 2 (■), 3 (▲), 4 (▼), 5 (◆) and 6 (★).

traditional percolation contexts). To this aim, we focus on the evolution with m of the size of the giant connected component, $S(m)$, as well as the evolution for the number of components $n_c(m)$, and that of average path length of the giant component, $L(m)$. In Figure 5, we show these evolutions for each of the 6 SES for the cases of Bogotá (top) and Medellín (bottom).

The adiabatic evolution of the giant component $S(m)$ shows that both cities behave in a similar way so that the different evolution $S(m)$ for the SES follows the same hierarchy. In particular, SES 2 and 3 reach to cover almost all the urban mobility zones of the cities. On the other hand, the coverage of SES 6 in both cities and also 5 in Bogotá are well below the 50% of the zones. The main difference between the two cities shows up by looking at the rate $S(m)$ increases. While for Medellín the rate of change is very small for all the SES, in Bogotá, SES 1, 2 and 3 need to merge at least two different transportation layers in order to achieve the 80% of their corresponding coverage. This result is the fingerprint of the segregation of these poor SES observed combined with the effect of the smaller sample size of the Bogotá survey (as compared to that of Medellín) that makes difficult to capture weak connections between urban mobility zones. These fact seems to affect more poor SES due again to their spatial segregation.

The evolution of the number of components $n_c(m)$, and the average path length $L(m)$ in the city of Bogotá further confirms the effects of the segregation of SES 1, 2 and 3. As observed, the initial ($m = 1$) values of both n_c and L are extremely large and they need to merge at least two transportation modes to reach small values of n_c and L . This is not the case for SES 4, 5 and 6 for which the evolution is far more

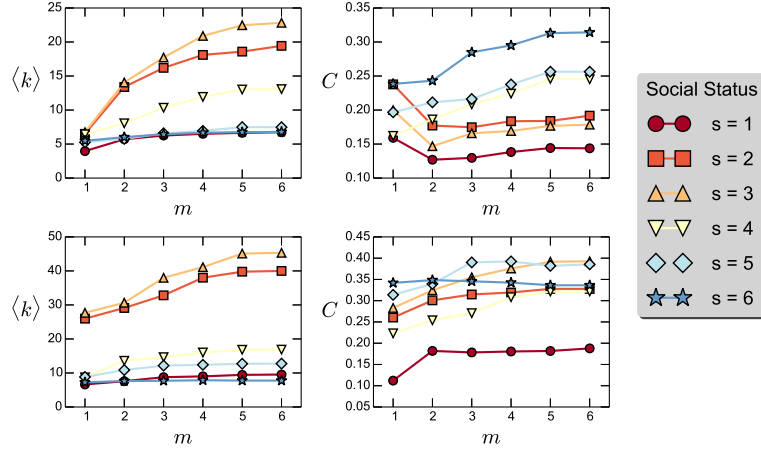


Fig. 6 Adiabatic evolution of the number of different destinations reached from a place, $\langle k \rangle$ and clustering coefficient, C as a function of the number of layers merged together, m . Left panels display the evolution for the average connectivity $\langle k \rangle(m)$, while those in the right show that of the clustering coefficient $C(m)$. The top panels are for the city of Bogotá, while those in the bottom refer to Medellín. Each curve corresponds to a different SES. Namely: 1 (●), 2 (■), 3 (▲), 4 (▼), 5 (◆) and 6 (★).

smooth. Concerning the final values of L in the city of Bogotá, it is remarkable the large steady value reached by SES 1 as compared to the rest of the population. Thus, even if they can cover a large number of zones the trips connecting them associate in a rather linear way, thus not displaying shortcuts. In its turn, the situation in Medellín concerning the evolution of $n_c(m)$ is not pretty much like to that of Bogotá. In fact, in this city the locations of the usual destinations appear to be very clustered, leading to have a system composed of only one component even for $m = 1$ for most of the SES. The evolution of $L(m)$ instead is more interesting. As in the case of Bogotá, L decreases with m although in a smoother way [as occurred for the evolution of $S(m)$]. Again, it is worth to notice how, as in the case of Bogotá, SES 1 displays a different behavior from the rest of strata.

Summarizing, both the behavior of $S(m)$ and $L(m)$ point out that in both Bogotá and (more clearly) Medellín the six SES can be regrouped into three mobility compartments related with their wealth. Namely: low (SES 1), mid-low (SES 2 and 3), and mid-high (SES 4, 5 and 6) compartments.

In Figure 6, we confirm the above compartmentalization by monitoring the evolution for the number of different trips from/to each urban area (here represented as the average degree $\langle k \rangle$ of the nodes) and the role of the various transportation modes in the triadic closure phenomenon, here studied via the clustering coefficient C . The evolution $\langle k \rangle(m)$ reveals two clearly distinct behaviors. First, for SES 2 and 3 (also 4 in the city of Bogotá) incorporating transportation modes implies to increase the number of origins/destinations, pointing out the genuine multimodal character of these individuals who assign different transportation modes depending on the trip to

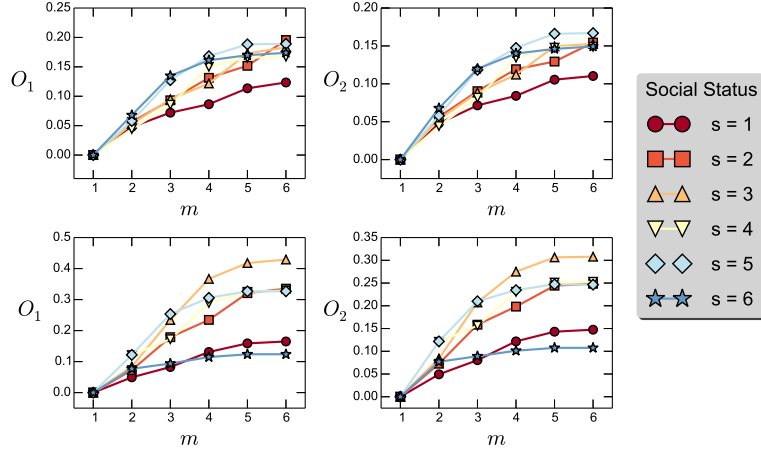


Fig. 7 Adiabatic evolution of the overlap as a function of the number of layers merged together, m . Left panels display the evolution for first definition of overlap, O_1 , Equation 5, while those in the right show the second one, O_2 , Equation 6. The top panels are for the city of Bogotá, while those in the bottom refer to Medellín. Each curve corresponds to a different SES. Namely: 1 (●), 2 (■), 3 (▲), 4 (▼), 5 (◆) and 6 (★).

be performed. On the other hand, for the rest of the SES there is almost no evolution. However, when looking back to the evolution of $S(m)$ in Figure 5, it is easy to notice that the almost steady behavior of $\langle k \rangle(m)$ for these strata has different roots. While individuals belonging to SES 5 and 6 move from/to a limited amount of different places [as displayed by the small values of $S(m)$] using few transportation modes, due to the aforementioned selection mechanism, SES 1 displays a large coverage. Thus, for SES 1, the addition of a new transportation layer is mostly devoted to join pairs of disconnected nodes, and thus not used to increment the communication power of zones for which a trip already exists.

The particular way of evolution with m displayed by SES 1 is also related to the large resulting networks [as displayed by $L(m)$ in Figure 5] and further confirmed by looking at the evolution of the clustering coefficient, $C(m)$. In both cities the values displayed by SES 1 are the smallest of the population and it does not show any significant change when increasing m . At variance, SES 5 and 6 display the largest values for the clustering in both cities, thus confirming again that, in this cases individuals cover a limited and rather fixed number of zones, thus favoring the formation of triadic paths in the aggregated graph.

Finally, in Figure 7 we present the evolution for two measures of the overlap proposed above (see Equations 5 and 6). Interestingly, for the two cities the two measures present almost the same evolution in terms of the relative growth between the different SES. Considering the definitions of O_1 (that takes into account the total amount of degeneration of the links) and O_2 (that only count once the redundant links regardless of the number of times they are repeated) it is clear that $O_1 \geq O_2$. However, the similar trends observed and the small difference in the values attained

by O_1 and O_2 point out that all of the SES do not tend to accumulate more than two overlapping links. Concerning the differences in the increase rates of O_1 and O_2 between SES we observe that, in both cities, individuals belonging to SES 1 tend to avoid overlapping, in agreement with the way (discussed before) SES 1 tend to increase the size of the giant component. Importantly, for the city of Bogotá the trends observed in both $O_1(m)$ and $O_2(m)$ seem to reproduce the three mobility compartments discussed above (1, 2 – 3 and 4 – 5 – 6) for both cities. However, the results for the city of Medellín are completely in disagreement with these compartments since, for instance, SES 6 display small overlapping tendency (being similar to that of SES 1) in contrast to the large tendency of SES 4 and 5.

5 Conclusions

We have presented a dataset about the human mobility in urban areas with two ingredients of utmost interest: the information about the multimodal nature of the trips, and the *socioeconomic status* of the individuals. The first ingredient has allowed us to tackle the analysis of the mobility patterns using the multiplex framework, which has attracted many attention lately. On the other hand, the information about SES provides with a novel ingredient that has been ignored up to date in the studies about human mobility. Exploiting these two ingredients, the aim of this manuscript has been to describe how the different socioeconomic compartments make use and combine the different transportation layers.

We have analyzed the mobility multiplex networks of each SES by studying how different structural descriptors evolve as network layers (the transportation modes) are merged. This procedure, called adiabatic projection, starts from the network of the trips performed by means of the most used transportation mode and subsequently adds the layers corresponding to the other means in descending order of usage.

The main result of our work is the classification of the 6 SES into three compartments according to their behavior. Namely, in a first group we have SES 1 and 2, the poorest ones, whose behavior is characterized by the segregation, *i.e.*, the usage of few and cheap transportation modes to cover a large fraction of the urban areas in a rather sparse way. The second compartment is composed of SES 3 and 4, having a genuine multimodal pattern and covering almost the total number of urban zones. Finally, the elite compartment composed of SES 5 and 6, is characterized by a selection of costly modes for performing the trips that, in their turn, display a very small coverage in terms of the urban areas reached although the connectivity within these zones turns to be rather dense.

The unveiled differences in the organization of the mobility multiplex networks according to SES demands the inclusion of this novel ingredient in the studies about human mobility and intrinsically related processes. As an example, it would be of interest to incorporate the presence of socioeconomic differences when studying the development of contagion processes in urban areas. We hope that our work will motivate more studies in this direction.

Acknowledgements We acknowledge financial support from the European Commission through FET IP projects MULTIPLEX (Grant No. 317532) and PLEXMATH (Grant No. 317614), from the Spanish MINECO under projects FIS2011-25167 and FIS2012-38266-C02-01, from the Comunidad de Aragón (Grupo FENOL), and from the Universidad Nacional de Colombia under grants HERMES 19010 and HERMES 16007. JGG is supported by the Spanish MINECO through the Ramón y Cajal program. AC acknowledge the financial support of SNSF through the project CRSII2_147609. We thank *Area Metropolitana del Valle de Aburrá*, in Medellín, and *Secretaría Distrital de Movilidad*, in Bogotá, for the Origin-Destination Surveys Datasets.

References

1. Stouffer, S. A.: Intervening opportunities: a theory relating mobility and distance. *Am. Sociol. Rev.* **5**, 845–867 (1940)
2. Zipf, G. K.: The $P_1 P_2 / D$ hypothesis: on the intercity movement of persons. *Am. Sociol. Rev.* **11**, 677–686 (1946)
3. Erlander, S., Stewart, N.: The Gravity Model in Transportation Analysis: Theory and Extensions. VSP (1990)
4. Batty, M.: The size, scale, and shape of cities. *Science*, **319**, 769–771 (2008)
5. Porta, S., Latora, V., Wang, F., Rueda, S., Strano, E., Scellato, S., Cardillo, A., Belli, E., Cárdenas, F., Cormenzana, B., Latora, L.: Street Centrality and Location of Economic Activities in Barcelona. *Urban Studies*, **49**, 1471–1488 (2011)
6. Porta, S., Latora, V., Wang, F., Strano, E., Cardillo, A., Scellato, S., Iacoviello, V., Messori, R.: Street centrality and densities of retail and services in Bologna, Italy. *Environment and Planning B: Planning and Design*, **36**, 450–465 (2009)
7. Guimerá, R., Mossa, S., Turtschi, A., Amaral, L.A.N.: The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proc. Nat. Acad. Sci. USA*, **102**, 7794–7799 (2005).
8. Asgari, F., Gauthier, V., Becker, M.: A survey on Human Mobility and its applications. *arXiv:1307.0814* (2013)
9. Yan, X.-Y., Han, X.-P., Wang, B.-H., Zhou, T.: Diversity of individual mobility patterns and emergence of aggregated scaling laws. *Scientific Reports*, **3**, 2678 (2013)
10. González, M. C., Hidalgo, C. a, Barabási, A.-L.: Understanding individual human mobility patterns. *Nature*, **453**, 779–782 (2008)
11. Wang, P., Hunter, T., Bayen, A. M., Schechtner, K., González, M. C.: Understanding Road Usage Patterns in Urban Areas. *Scientific Reports*, **2**, 1001 (2012)
12. Roth, C., Kang, S. M., Batty, M., Barthélemy, M.: Structure of urban movements: polycentric activity and entangled hierarchical flows. *PLoS ONE*, **6**, e15923 (2011)
13. Helbing, D., Buzna, L., Johansson, A., Werner, T.: Self-Organized Pedestrian Crowd Dynamics: Experiments, Simulations, and Design Solutions. *Transp. Sci.*, **39**, 1–24, (2005).
14. Bazzani, A., Giorgini, B., Rambaldi, S., Turchetti, G.: ComplexCity: modeling urban mobility. *Advances in complex system (ACS)*, **10**, 255–270 (2007).
15. Song, C., Koren, T., Wang, P., Barabási, A.-L.: Modelling the scaling properties of human mobility. *Nat. Phys.*, **6**, 818–823 (2010).
16. Simini, F., González, M. C., Maritan, A., Barabási, A.-L.: A universal model for mobility and migration patterns. *Nature*, **484**, 96–100 (2012).
17. Eubank, S., Guclu, H., Kumar, V., Marathe, M.: Modelling disease outbreaks in realistic urban social networks. *Nature*, **429**, 180–184, (2004)
18. Colizza, V., Barrat, A., Barthélemy, M., Vespignani, A.: The role of the airline transportation network in the prediction and predictability of global epidemics. *Proc. Nat. Acad. Sci. USA*, **103**, 2015–2020 (2006).
19. Kleinberg, J.: Computing: the wireless epidemic. *Nature*, **449**, 287 (2007).

20. Balcan, D., Colizza, V., Gonçalves, B., Hu, H., Ramasco, J. J., Vespignani, A.: Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Nat. Acad. Sci. USA*, **106**, 21484 (2009).
21. Tizzoni, M., Bajardi, P., Poletto, C., Ramasco, J. J., Balcan, D., Gonçalves, B., Perra, N., Colizza, V., Vespignani, A.: Real-time numerical forecast of global epidemic spreading: case study of 2009 A/H1N1pdm. *BMC Medicine*, **10**, 165 (2012).
22. Poletto, C., Tizzoni, M., Colizza, V.: Human mobility and time spent at destination: impact on spatial epidemic spreading. *J. Theor. Bio.*, **338**, 41–58 (2013).
23. Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47 (2002).
24. Newman, M.E.J.: The structure and function of complex networks. *SIAM Review* **45**, 167–256 (2003).
25. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.: Complex networks: Structure and dynamics. *Physics Reports*, **424**, 175–308 (2006).
26. Barthélemy, M.: Spatial networks. *Physics Reports*, **499**, 1-101 (2011).
27. Zaltz Austwick, M., O'Brien, O., Strano, E., Viana, M.: The Structure of Spatial Networks and Communities in Bicycle Sharing Systems. *PLoS ONE* **8** e74685 (2013).
28. De Domenico, M., Solé-Ribalta, A., Cozzo, E., Kivelä, M., Moreno, Y., Porter, M. A., Arenas, A.: Mathematical Formulation of Multilayer Networks. *Phys. Rev. X*, **3**, 041022 (2013).
29. Boccaletti, S., Bianconi, G., Criado, R., Del Genio, C.I., Gómez-Gardeñes, J., Romance, M., Sendiña-Nadal, I., Wang, Z., Zanin, M.: The structure and dynamics of Multilayer Networks. *Physics Reports*, in press (2014).
30. Kivelä, M., Arenas, A., Barthélemy, M., Gleeson, J. P., Moreno, Y., Porter, M. A.: Multilayer Networks. *Journal of Complex Networks*, in press (2014).
31. De Domenico, M., Solé-Ribalta, A., Gómez, S., Arenas, A.: Navigability of interconnected networks under random failures. *Proc. Nat. Acad. Sci. USA*, in press (2014).
32. Kurant, M., Thiran, P.: Layered Complex Networks. *Phys. Rev. Lett.*, **96**, 138701 (2006).
33. Cardillo, A., Zanin, M., Gómez-Gardeñes, J., Romance, M., García del Amo, A. J., Boccaletti, S.: Modeling the multi-layer nature of the European Air Transport Network: Resilience and passengers re-scheduling under random failures. *Eur. Phys. J. Special Topics*, **215**, 23–33 (2013).
34. Cardillo, A., Gómez-Gardeñes, J., Zanin, M., Romance, M., Papo, D., Del Pozo, F., Boccaletti, S.: Emergence of network features from multiplexity. *Scientific Reports*, **3**, 1344 (2013).
35. Barigozzi, M., Fagiolo, G., Garlaschelli, D.: Multinetwork of international trade: A commodity-specific analysis. *Phys. Rev. E*, **81**, 046104 (2010).
36. Bianconi, G.: Statistical mechanics of multiplex networks: Entropy and overlap. *Phys. Rev. E*, **87**, 062806 (2013).
37. Kapferer, B.: Norms and the manipulation of relationships in a work context. In J. C. Mitchell, editor, *Social Networks in Urban Situations: Analyses of Personal Relationships in Central African Towns*. Manchester University Press, (1969).
38. Parshani, R., Rozenblat, C., Ietri, D., Ducruet, C., Havlin, S.: Inter-similarity between coupled networks. *Europhys. Lett.*, **92**, 68002 (2010).
39. Battiston, F., Nicosia, V., Latora, V.: Structural measures for multiplex networks. *Phys. Rev. E*, **89**, 032804 (2014).
40. Secretaría Distrital de Movilidad. (2011). Informe de indicadores Encuesta de Movilidad de Bogotá 2011. Bogotá: Unión Temporal Steer Davies & Gleave Limited - Centro Nacional de Consultoría.
Retrieved from <http://www.movilidadbogota.gov.co/?pag=1246>
41. AREA Metropolitana del Valle de Aburrá. (2006). Capítulo 2: Diagnóstico. Formulación del Plan Maestro de Movilidad para la Región Metropolitana del Valle de Aburrá. Informe Final (pp. 21-72).
Retrieved from: <http://www.areadigital.gov.co/Movilidad/Documents/Plan%20Maestro%20de%20Movilidad.pdf>